Heart rate variability reproducibility during exercise

To cite this article: Melitta A McNarry and Michael J Lewis 2012 Physiol. Meas. 33 1123

View the article online for updates and enhancements.
Heart rate variability reproducibility during exercise

Melitta A McNarry and Michael J Lewis

College of Engineering, Swansea University, Wales, UK

E-mail: m.mcnarry@swansea.ac.uk

Received 30 April 2012, accepted for publication 7 June 2012
Published 27 June 2012
Online at stacks.iop.org/PM/33/1123

Abstract
The use of heart rate variability (HRV) parameters during exercise is not supported by appropriate reliability studies. In 80 healthy adults, ECG was recorded during three 6 min bouts of exercise, separated by 6 min of unloaded cycling. Two bouts were at a moderate intensity while the final bout was at a heavy exercise intensity. This protocol was repeated under the same conditions on three occasions, with a controlled start time (pre-determined at the first visit). Standard time and frequency domain indices of HRV were derived. Reliability was assessed by Bland–Altman plots, 95% limits of agreement and intraclass correlation coefficients (ICC). The sample size required to detect a mean difference ≥30% of the between-subject standard deviation was also estimated. There was no systematic change between days. All HRV parameters demonstrated a high degree of reproducibility during baseline (ICC range: 0.58–0.75), moderate (ICC: 0.58–0.85) and heavy intensity exercise (ICC range: 0.40–0.76). The reproducibility was slightly diminished during heavy intensity exercise relative to both unloaded baseline cycling and moderate exercise. This study indicates that HRV parameters can be reliably determined during exercise, and it underlines the importance of standardizing exercise intensity with regard to fitness levels if HRV is to be reliably determined.

Keywords: exercise intensity, ECG, coefficient of variation, intraclass correlation coefficient, sample size

Introduction

Heart rate variability (HRV) has been studied extensively across a myriad of academic areas as it provides a non-invasive insight into the function of the autonomic nervous system (Akselrod et al 1981, TaskForce 1996). Epidemiological studies have utilized HRV to characterize...
differences in autonomic function according to age (Liao et al. 1995, Odemuyiwa et al. 1992), aerobic fitness level (Aubert et al. 2001, Melanson 2000, Carter et al. 2003) and between genders (Stein et al. 1997, Umetani et al. 1998). Furthermore, HRV has been shown to be sensitive to pathological conditions, such as heart failure (e.g. Stein et al. 1995, Casolo et al. 1989), hypertension (Konrady et al. 2001) and diabetes (e.g. Burger et al. 1997, Nolan et al. 1996), and to be prognostically significant (Kleiger et al. 1987, Bigger et al. 1992, Tsuji et al. 1994, 1996). Fundamental to interpreting such between-group differences in HRV and so interpreting their clinical or epidemiological significance is verification of the reliability and reproducibility of these variables.

A substantial body of evidence is available regarding the reliability of HRV parameters during both 24 h recordings (e.g. Zuanetti et al. 1991, Ziegler et al. 1999, Bigger et al. 1992, Pitzalis et al. 1996) and short-term (e.g. 2–15 min) recordings under stationary, stable conditions (e.g. Schroeder et al. 2004, Lord et al. 2001, Sandercock et al. 2004, Pinna et al. 2007). However, although HRV assessment during a resting steady state is attractive to minimize data artefacts, the environmental validity of such a state is limited.

Given the profound influence of exercise on HRV (Al-Ani et al. 1996, Gladwell et al. 2005, Gladwell and Coote 2002, Raven et al. 2006, Ogoh et al. 2002), it might be hypothesized that HRV reproducibility could be significantly influenced under exercising conditions. Few studies have investigated the reproducibility of HRV during exercise and inter-study comparisons are largely prohibited by the use of different exercise protocols and HRV parameters. Tulpo et al. (1998) reported the reliability of SD1 and HF to increase with increasing exercise intensity whilst Gujit et al. (2007) demonstrated high intraclass correlation coefficients (ICC) for SDNN and RMSSD during 15 min cycling at 50 W. These previously investigated HRV parameters are widely recognized as surrogates of vagal modulation. A limitation of these earlier studies is the lack of standardization of exercise intensity between participants; comparing participants at the same absolute intensity (e.g. 50 W) fails to account for heterogeneous fitness levels which could result in participants differing considerably in the relative difficulty of this absolute work rate. Such differences might introduce non-physiological variation (e.g. resulting from mechanical or movement-induced artefact), thereby influencing conclusions regarding the reliability of HRV during exercise. Indeed, the potential importance of exercise intensity is suggested by the results of Tulpo et al. (1998).

The purpose of this study was therefore to investigate HRV reproducibility during physical exercise. To explore the influence of exercise intensity on HRV reliability during exercise, two relative exercise intensities were studied: moderate and heavy intensity exercise. We sought to quantify reliability for both time and frequency domain indices of HRV during exercise conditions. HRV reproducibility was investigated within and between participants and within/between exercise sessions. We hypothesized 1) that HRV reproducibility would be greater during heavy compared with moderate intensity exercise, and 2) that time domain HRV variables would be more reproducible than frequency domain variables.

**Methods**

**Participants**

Eighty adults (mean ± S.D. age 33 ± 11 years; body mass 74.3 ± 12.6 kg; height 1.7 ± 0.1 m; BMI 24 ± 3 kg·m⁻²; 49 male) volunteered for the study. The participants were all recreationally active, but not highly trained. Prior to testing, participants were informed of the protocol and risks and gave written consent to participate. All procedures were approved by the local ethics committee and were conducted in accordance with the Declaration of Helsinki.
Participants were asked to arrive at the laboratory in a rested and fully hydrated state, at least 2 h postprandial and to avoid strenuous exercise in the 24 h preceding each testing session. Participants were also asked to refrain from caffeine and alcohol 6 and 24 h before each test, respectively. All tests were performed at the same time of day (±2 h).

**Experimental design**

Participants were required to visit the laboratory on four occasions, separated by at least 24 h of recovery. Participants initially completed a ramp incremental exercise test for determination of $\dot{V}O_{2\text{peak}}$ and the gas exchange threshold (GET). On each of the three subsequent visits, participants completed two bouts of moderate intensity exercise (Mod: at a work rate calculated to elicit 70% of the GET) followed by a bout of heavy intensity exercise (Hvy: at a work rate calculated to elicit a $\dot{V}O_2$ equal to the GET plus 30% of the difference between the GET and peak $\dot{V}O_2$, i.e. $\delta 30\%$). All exercise testing was conducted using an electronically braked cycle ergometer (Lode Excalibur, Groningen, Netherlands).

**Incremental test**

Initially, participants completed 3 min of baseline cycling at 0 W, after which the work rate was increased at a rate of 20–30 W min$^{-1}$ until the limit of tolerance. The participants were asked to maintain a cadence of 70–80 rpm. Breath-by-breath pulmonary gas-exchange data were collected continuously during the incremental tests and averaged over consecutive 5 s periods (Oxycon Pro, Jaeger, Germany). The $\dot{V}O_{2\text{peak}}$ was taken as the highest 10 s average value attained before the subject’s volitional exhaustion in the test. The GET was determined from a cluster of measurements, including (1) the first disproportionate increase in CO$_2$ production ($\dot{V}O_2$) from visual inspection of individual plots of $\dot{V}O_2$ versus $\dot{V}O_2$; (2) an increase in expired ventilation $\dot{V}_E/\dot{V}O_2$ with no increase in $\dot{V}_E/\dot{V}CO_2$; and (3) an increase in end-tidal O$_2$ tension with no fall in end-tidal CO$_2$ tension. The work rates that would require 70% of the GET (moderate exercise) and $\Delta 30\%$ were subsequently determined, accounting for the mean response time for $\dot{V}O_2$ during ramp exercise (i.e. two thirds of the ramp rate was deducted from the work rate at the GET and peak (Whipp et al 1981)).

**Step exercise tests**

To investigate the reliability and reproducibility of HRV parameters during exercise, participants completed a series of ‘step’ tests. The protocol, which was repeated three times on separate days, comprised of two moderate intensity and one heavy intensity cycle transition, each of 6 min duration. Each transition was preceded by 6 min of baseline pedalling at 0 W followed by an abrupt transition to the target work rate. Therefore, all participants performed a total of six bouts of moderate-intensity exercise and three bouts of heavy intensity exercise. Throughout each of the exercise transitions, participants were asked to maintain a pedal cadence of 70 ± 5 rpm. Recent studies have highlighted the potential influence of pedal cadence on HRV, this influence increasing with increasing work rate (Blain et al 2009). To minimize the potential for this to confound the present results, the pedal cadence was strictly controlled and a constant, relative work rate was used.

**Measurements**

Throughout all exercise tests, participants wore a facemask and breathed through a low dead space (90 ml), low resistance (0.75 mmHg l$^{-1}$ s$^{-1}$ at 15 l s$^{-1}$) impeller turbine assembly (Jaeger
Triple V, Hoechberg, Germany). The inspired and expired gas volumes and gas concentration signals were continuously sampled at 100 Hz, the latter using paramagnetic (O2) and infrared (CO2) analysers (Jaeger Oxycon Pro, Hoechberg, Germany) via a capillary line connected to the mouthpiece. These analysers were calibrated before each test with gases of known concentrations, and the turbine volume transducer was calibrated using a 3 l syringe (Hans Rudolph, Kansas City, MO). The volume and concentration signals were time aligned by accounting for the delay in capillary gas transit and analyser rise time relative to the volume signal.

A Reynolds Lifecard CF digital Holter recorder (Spacelabs Medical Ltd, Hertford, UK) was used to record a three-lead ECG continuously throughout the tests. The ECG leads were positioned in the modified V5, CC5, modified V5R electrode configuration. This system provided ECG data with a sample accuracy of 2.5 μV (magnitude of least significant bit; 12-bit resolution) and 1024 Hz sampling frequency.

**ECG analysis**

ECG recordings were analysed using a Reynolds Pathfinder digital analyser (Spacelabs Medical Ltd, UK). Beat-to-beat cardiac interval (RR) values were automatically measured for each sinus beat and exported for further analysis using the Reynolds Research Tools software (Spacelabs Medical Ltd, UK). All ECG data used for subsequent analysis in this study were free of any form of morphologically abnormal beat, and this was verified by both the Holter system and by visual inspection.

**Heart rate variability derivation**

HRV variables were quantified in the time domain (RMSSD: square root of the mean of the sum of the squares of differences between adjacent RR intervals; SDNN: standard deviation of all RR intervals) and the frequency domain (total power: 0.017–0.4 Hz, low frequency (LF) power: 0.04–0.15 Hz, high frequency (HF) power: 0.15–0.4 Hz) according to the Task Force guidelines on HRV (TaskForce 1996). Additionally, to account for the higher respiratory frequencies evident during exercise and their influence on HRV (Lewis et al 2007, Bailon et al 2010), we also assessed HF and TP using extended frequency bandwidths of 0.15–0.87 Hz and 0.017–0.87 Hz, respectively. The upper limit of these bandwidths was selected according to the highest respiratory frequency observed in this study (52 breaths · min⁻¹). Prior to the frequency domain analysis procedures RR interval data were re-sampled using a sampling frequency of 3 Hz; this is closest integer to the mean maximal cardiac cycle frequency (155 bpm being equivalent to ~2.6 Hz), and is sufficient also to avoid potential aliasing from any cycle cadence-related component (where cadence frequency here was ~1.3 Hz). The RR interval data were then linearly de-trended and Hanning windowed in consecutive 1 min segments; the power spectral density of each segment was then calculated using the Welch periodogram method, using short-term Fourier transformation and a 50% overlap between adjacent segments.

Baseline HRV parameters were calculated as the mean value from the middle 3 min of unloaded cycling (i.e. minutes 3–5) and steady state exercise values as the mean value between 3.5 and 5.5 min after the onset of an increased external work rate.

**Statistical analysis**

Initially, data were assessed for normality by the Shapiro–Wilks test; those revealed as skewed were log-transformed prior to subsequent analyses. The distribution of the measurements
Table 1. HRV variables at baseline and during steady state moderate and heavy intensity exercise.

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Moderate</th>
<th>Heavy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Mean Δ</td>
<td></td>
</tr>
<tr>
<td>SDNN (ms)</td>
<td>35.6 ± 12.8</td>
<td>0.02</td>
<td>28.3 ± 12.0</td>
</tr>
<tr>
<td>Ln SDNN (ln ms)</td>
<td>3.5 ± 0.4</td>
<td>0.00</td>
<td>3.2 ± 0.4</td>
</tr>
<tr>
<td>RMSSD (ms)</td>
<td>24.1 ± 10.2</td>
<td>−0.44</td>
<td>19.3 ± 9.7</td>
</tr>
<tr>
<td>Ln RMSSD (ln ms)</td>
<td>3.1 ± 0.4</td>
<td>−0.05</td>
<td>2.8 ± 0.5</td>
</tr>
<tr>
<td>LF (ms²)</td>
<td>6.9 ± 3.0</td>
<td>0.04</td>
<td>5.0 ± 2.9</td>
</tr>
<tr>
<td>Ln LF (ln ms²)</td>
<td>1.8 ± 0.5</td>
<td>0.00</td>
<td>1.4 ± 0.6</td>
</tr>
<tr>
<td>HF0.4 (ms²)</td>
<td>5.1 ± 2.4</td>
<td>−0.19</td>
<td>3.7 ± 2.1</td>
</tr>
<tr>
<td>Ln HF0.4 (ln ms²)</td>
<td>1.5 ± 0.5</td>
<td>−0.06</td>
<td>1.1 ± 0.6</td>
</tr>
<tr>
<td>TP (ms²)</td>
<td>147.4 ± 23.9</td>
<td>−2.36</td>
<td>128.7 ± 27.0</td>
</tr>
<tr>
<td>Ln TP (ln ms²)</td>
<td>5.0 ± 0.2</td>
<td>−0.10</td>
<td>4.8 ± 0.2</td>
</tr>
<tr>
<td>LFnu</td>
<td>0.4 ± 0.1</td>
<td>0.01</td>
<td>0.4 ± 0.1</td>
</tr>
<tr>
<td>Ln LFnu</td>
<td>−0.8 ± 0.2</td>
<td>0.05</td>
<td>0.9 ± 0.3</td>
</tr>
<tr>
<td>HF0.87 (ms²)</td>
<td>8.0 ± 3.2</td>
<td>−0.21</td>
<td>6.6 ± 3.6</td>
</tr>
<tr>
<td>Ln HF0.87 (ln ms²)</td>
<td>2.0 ± 0.4</td>
<td>−0.06</td>
<td>1.8 ± 0.5</td>
</tr>
</tbody>
</table>

Mean ± S.D. N = 80; SDNN, standard deviation of all RR intervals; RMSSD, square root of the mean of the sum of the squares of differences between adjacent RR intervals; LF, low frequency power; HF0.4, high frequency power (0.15–0.4 Hz); TP, total power; LFnu, normalized low frequency power; HF0.87, high frequency power (0.15–0.87 Hz); Ln, natural log. All mean differences (Mean Δ) between tests were non-significant.

obtained during each of the tests, and the inter-test difference, was then examined to identify and remove outliers. Bland–Altman plots of the difference between paired measurements as a function of their mean were constructed to allow visualization of any systematic change between the tests and to assess the data for heteroscedasticity. The absence of any systematic between-test change was further verified by paired samples t-tests.

The standard error of measurement (SEM) was estimated from the results of an ANOVA as the square root of the within-subject mean square. The 95% limits of random variation were subsequently derived from the SEM values, with the limits back-transformed for those HRV variables initially log-transformed. ICC were also determined from the mean square values of the ANOVA.

As the reliability of measurements influences the sample size required for an experiment, we estimated the sample size required to detect a relevant change in the mean of the HRV parameters after a treatment (test–retest design) (Lehr 1992). A change of ≥30% of the between participant mean was considered to be ‘relevant’ and was used in a two-sided test with a significance level of 5% and power of 0.8.

All analyses were conducted using the PASW Statistics package version 18 (SPSS, Chicago, IL). Statistical significance was accepted as P < 0.05.

Results

Descriptive statistics for the HRV variables at baseline and during both exercise intensities are shown in table 1. All the HRV variables demonstrated a significantly skewed distribution, according to the Shapiro–Wilks test (P < 0.05), at both baseline and steady state exercise and during both exercise intensities. Bland–Altman plots revealed a symmetrical distribution of points around the zero line, indicating the absence of a systematic change as a function of the mean (figure 1(A)). Such an absence was further verified by non-significant paired samples t-tests (P > 0.05). The Bland–Altman plots did, however, reveal a trend where the scatter
Figure 1. Representative Bland–Altman plots of the difference between two repeat measurements under the same conditions (moderate). The data points were symmetrically distributed about the zero line but the degree of scatter around the line tended to increase in (A), indicating heteroscedasticity. Log-transformation was successful in ameliorating this trend and in eliciting homoscedasticity (B).

(difference values) increased with increasing mean values, suggesting heteroscedasticity in the data. Log-transformation of these variables elicited both a normal distribution (no skew) and homoscedasticity (figure 1(B)).

Reliability indices at baseline and during moderate and heavy intensity exercise are shown in tables 2 and 3. A good-to-excellent reliability was evident for all HRV variables during all conditions; heavy intensity exercise was associated with the lowest reliability with moderate intensity exercise demonstrating the greatest reproducibility for the majority of HRV variables.

According to the sample size formula proposed by Lehr (1992), these results suggest that a sample size of 174 participants would be required to detect a pre–post treatment difference of $\geq 30\%$ of the between-participant standard deviation for each HRV parameter.

**Discussion**

This study demonstrated ‘good to excellent’ reproducibility of HRV parameters, both at rest and during exercise. The ICCs calculated here during moderate intensity exercise are
Table 2. Reliability indices of HRV parameters at baseline.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SEM</th>
<th>95% CI (lower, upper)</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDNN (ms)</td>
<td>0.59</td>
<td>0.54, 0.64</td>
<td>0.66</td>
</tr>
<tr>
<td>Ln SDNN (ln ms)</td>
<td>0.58</td>
<td>0.53, 0.63</td>
<td>0.67</td>
</tr>
<tr>
<td>RMSSD (ms)</td>
<td>0.56</td>
<td>0.51, 0.61</td>
<td>0.69</td>
</tr>
<tr>
<td>Ln RMSSD (ln ms)</td>
<td>0.52</td>
<td>0.48, 0.57</td>
<td>0.73</td>
</tr>
<tr>
<td>LF (ms²)</td>
<td>0.59</td>
<td>0.54, 0.64</td>
<td>0.66</td>
</tr>
<tr>
<td>Ln LF (ln ms²)</td>
<td>0.59</td>
<td>0.54, 0.65</td>
<td>0.66</td>
</tr>
<tr>
<td>HF₀.₄ (ms²)</td>
<td>0.63</td>
<td>0.58, 0.69</td>
<td>0.61</td>
</tr>
<tr>
<td>Ln HF₀.₄ (ln ms²)</td>
<td>0.58</td>
<td>0.53, 0.63</td>
<td>0.67</td>
</tr>
<tr>
<td>TP (ms²)</td>
<td>0.53</td>
<td>0.49, 0.59</td>
<td>0.72</td>
</tr>
<tr>
<td>Ln TP (ln ms²)</td>
<td>0.51</td>
<td>0.47, 0.56</td>
<td>0.75</td>
</tr>
<tr>
<td>LFn</td>
<td>0.65</td>
<td>0.60, 0.71</td>
<td>0.58</td>
</tr>
<tr>
<td>Ln LFn</td>
<td>0.62</td>
<td>0.57, 0.68</td>
<td>0.62</td>
</tr>
<tr>
<td>HF₀.₈₇ (ms²)</td>
<td>0.60</td>
<td>0.55, 0.66</td>
<td>0.65</td>
</tr>
<tr>
<td>Ln HF₀.₈₇ (ln ms²)</td>
<td>0.57</td>
<td>0.53, 0.63</td>
<td>0.68</td>
</tr>
</tbody>
</table>

N = 80; SEM, standard error of measurement; CI, confidence interval; ICC, intraclass correlation coefficient; SDNN, standard deviation of all RR intervals; RMSSD, square root of the mean of the sum of the squares of differences between adjacent RR intervals; LF, low frequency power; HF₀.₄, high frequency power (0.15–0.4 Hz); TP, total power; LFn, normalized low frequency power; HF₀.₈₇, high frequency power (0.15–0.87 Hz); Ln, natural log.

Table 3. Reliability indices of HRV parameters during moderate and heavy intensity exercise.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Moderate</th>
<th>SEM</th>
<th>95% CI (±)</th>
<th>ICC</th>
<th>Heavy</th>
<th>SEM</th>
<th>95% CI (±)</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDNN (ms)</td>
<td></td>
<td>0.49</td>
<td>0.45, 0.54</td>
<td>0.76</td>
<td>0.73</td>
<td>0.64, 0.85</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>Ln SDNN (ln ms)</td>
<td></td>
<td>0.51</td>
<td>0.47, 0.56</td>
<td>0.75</td>
<td>0.72</td>
<td>0.64, 0.83</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>RMSSD (ms)</td>
<td></td>
<td>0.53</td>
<td>0.49, 0.58</td>
<td>0.73</td>
<td>0.65</td>
<td>0.57, 0.75</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>Ln RMSSD (ln ms)</td>
<td></td>
<td>0.54</td>
<td>0.49, 0.59</td>
<td>0.72</td>
<td>0.61</td>
<td>0.54, 0.71</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>LF (ms²)</td>
<td></td>
<td>0.59</td>
<td>0.54, 0.65</td>
<td>0.66</td>
<td>0.79</td>
<td>0.70, 0.92</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Ln LF (ln ms²)</td>
<td></td>
<td>0.59</td>
<td>0.54, 0.64</td>
<td>0.66</td>
<td>0.69</td>
<td>0.61, 0.81</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>HF₀.₄ (ms²)</td>
<td></td>
<td>0.52</td>
<td>0.48, 0.57</td>
<td>0.74</td>
<td>0.70</td>
<td>0.62, 0.82</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>Ln HF₀.₄ (ln ms²)</td>
<td></td>
<td>0.51</td>
<td>0.47, 0.56</td>
<td>0.74</td>
<td>0.67</td>
<td>0.59, 0.78</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>TP (ms²)</td>
<td></td>
<td>0.39</td>
<td>0.36, 0.42</td>
<td>0.85</td>
<td>0.51</td>
<td>0.45, 0.60</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Ln TP (ln ms²)</td>
<td></td>
<td>0.39</td>
<td>0.36, 0.43</td>
<td>0.85</td>
<td>0.50</td>
<td>0.44, 0.58</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>LFn</td>
<td></td>
<td>0.65</td>
<td>0.60, 0.71</td>
<td>0.58</td>
<td>0.78</td>
<td>0.69, 0.81</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>Ln LFn</td>
<td></td>
<td>0.64</td>
<td>0.59, 0.70</td>
<td>0.59</td>
<td>0.73</td>
<td>0.65, 0.85</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>HF₀.₈₇ (ms²)</td>
<td></td>
<td>0.48</td>
<td>0.44, 0.53</td>
<td>0.77</td>
<td>0.73</td>
<td>0.65, 0.85</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>Ln HF₀.₈₇ (ln ms²)</td>
<td></td>
<td>0.54</td>
<td>0.50, 0.59</td>
<td>0.71</td>
<td>0.73</td>
<td>0.64, 0.84</td>
<td>0.48</td>
<td></td>
</tr>
</tbody>
</table>

N = 80; SEM, standard error of measurement; CI, confidence interval; ICC, intraclass correlation coefficient; SDNN, standard deviation of all RR intervals; RMSSD, square root of the mean of the sum of the squares of differences between adjacent RR intervals; LF, low frequency power; HF₀.₄, high frequency power (0.15–0.4 Hz); TP, total power; LFn, normalized low frequency power; HF₀.₈₇, high frequency power (0.15–0.87 Hz); Ln, natural log.

considerably higher than the values of 0.62 and 0.57 previously reported during exercise at 50 W for SDNN and RMSSD, respectively (Guijt et al. 2007). This higher reproducibility is likely attributable to our use of relative exercise intensities, and the higher number of repeat visits and larger sample size in our study. No ICCs for other HRV parameters are presently available for comparison. This is the first study to address the influence of exercise intensity on HRV reproducibility. Contrary to our hypothesis, we observed a diminished reproducibility (by 3–33%), for all HRV parameters during heavy intensity exercise compared with moderate
intensity exercise. This might have been anticipated and could be (at least partly) attributable to the detrimental influence of the more pronounced movements associated with heavier exercise and the consequently greater number of movement artefacts in the data. This finding should be borne in mind for future studies: attempts should be made to mitigate the influence of movement-related artefact on HRV at high workloads by using appropriate motion-limiting protocols.

The process of identifying outliers undertaken here (identification of points outside the upper and lower quartile by more than 1.5 times the interquartile range) has been advocated to minimize the influence of such movements. Moreover, when such procedures are followed, the present results indicate that HRV parameters can be reliably determined even during heavy intensity exercise, with an acceptable level of confidence. The lower reproducibility observed here during heavy intensity exercise might also be related to the tertiary influence of respiratory activity on the HRV parameters. The activity of the cardiac vagal nerve (reflected here principally in the values of RMSSD and HF) is modulated by respiration (Porges and Byrne 1992), with suggestions that there is central integration of some cardiovascular and respiratory processes (Eckberg 2009). Given the greater respiratory activity necessitated by heavy intensity exercise, this could be responsible to some extent for confounding the measurements of HRV during these conditions. Although both heart period fluctuation and respiratory activity are reproducible, their respective patterns and inter-relationship could be altered according to exercise intensity. As far as we are aware, there have been no reports of either the reproducibility of respiratory sinus arrhythmia during exercise or of the influence of exercise intensity on the relationship between RR period fluctuations and respiratory activity.

Interestingly, the HRV parameters were also either equally reliable or more reliable during moderate exercise than during unloaded baseline cycling. This finding, which agrees with that of Tulppo et al (1998) during incremental exercise, is likely a reflection of the profound influence of exercise on the ANS (Al-Ani et al 1996, Gladwell and Coote 2002, Gladwell et al 2005, Ogoh et al 2002, Raven et al 2006); exercise minimizes the influence of many potentially confounding factors such as a participant’s mood, alertness and mental activity (Pinna et al 2007). This finding also supports the notion that using an absolute work rate (exercise intensity) is inappropriate for the reliable determination of HRV parameters. Assigning an absolute work rate to a group of individuals of differing fitness or ‘work capacity’ would yield inconsistent HRV reliability since individuals would be working at different relative intensities. Reliability of HRV measures would thus differ between individuals and the mean reliability would be misrepresentative. Therefore we advise that relative work rates are assigned in HRV study protocols, especially when there is heterogeneity of fitness levels in the study group.

The present results illustrate that HRV variables are not subject to systematic change, even over many repeated trials. This may be pertinent in informing the debate regarding the mechanistic basis and physiological relevance of HRV, for example indicating that exercise ‘learning’ effects do not exert a substantial influence on the magnitude of HRV variables. Our observed between-day reliability agrees with some (Kowalewski and Urban 2004, Kleiger et al 1991, Bigger et al 1992, Tulppo et al 1998) but not all (Lord et al 2001, Pinna et al 2007) previous studies. Such ambiguity within the literature is, at least in part, attributable to inter-study differences in the time interval between tests (from one day to 24 months apart), the conditions of the test (e.g. supine, standing, spontaneous breathing, paced breathing), participant characteristics (e.g. healthy, cardiac transplant, young, older) and the specific statistical comparisons made (e.g. between- or within-participant assessments of day-to-day variability). Furthermore, Schroeder et al (2004) also found that the reproducibility of measurements increased with measurement duration, advocating a minimum of 5 min recordings for epidemiological studies. The present results extend those recommendations.
by suggesting that 2–2.5 min recordings provide reproducible measures of HRV during both moderate and heavy intensity exercise.

In contrast to our second hypothesis, there was no uniform difference in the repeatability between time-domain and frequency-domain HRV variables. It appears that both methods of HRV quantification yield similarly reproducible indices of HRV during exercise. We thought that the time domain parameters would demonstrate a greater reproducibility as these are suggested to be less influenced by non-stationarities in the data (Pagani et al 1988, Perandini et al 2009, Asmussen and Kristiansson 1968). However, it must be noted that quantification of HRV in the frequency-domain requires pre-processing to improve the stability of the RR time-series, and this might conceivably have improved the repeatability of the frequency-domain HRV measures. Nevertheless, such pre-processing is an essential procedure and so we believe our results are pragmatic.

Prior to log-transformation, we observed significant heteroscedasticity in the data, that is, a trend for the scatter of the data points to increase with increasing mean values. Recent studies have illustrated the dependence of HRV parameters on mean heart rate (Meste et al 2005, Chiu et al 2003, Bailo et al 2011), and it might be suggested that such dependence contributes to the heteroscedasticity observed here. However, this heteroscedasticity was observed for each constant, relative exercise intensity, and we therefore consider that there is minimal likelihood of this potential explanation.

Although not the purpose of this study, it is pertinent to note (as in many other studies) the discernible influence of exercise intensity on HRV; heavy intensity exercise is associated with a greater change than moderate intensity exercise. This supports the contention of a two-phase alteration in autonomic regulation during exercise, consisting of an initial withdrawal of vagal modulation and a more gradual secondary increase in sympathetic tone (Al-Ani et al 1996, Gladwell and Coote 2002, Gladwell et al 2005, Ogor et al 2002, Raven et al 2006). There is still much scope for HRV research during exercise: for example, it remains to be elucidated if the dramatic changes in HRV during heavy intensity exercise are influenced by factors such as age and the interaction between age and fitness. The present results should ameliorate any hitherto concerns regarding the reproducibility of HRV parameters during heavy intensity exercise, and its potential confounding effect on studies such as the above.

Limitations

A potential limitation of this study was the restriction of the data analysis to the use of parametric techniques, requiring log-transformation to meet the required assumptions of normality. This method was adopted in line with previous studies to allow inter-study comparisons to be made. Future studies might usefully address the influence of parametric versus non-parametric statistical methods of analysis. Another limitation of this study was the large heterogeneity in the participant ages, which ranged from 18 to 58 years. However, a similar range and standard deviation in participant ages has often been reported previously (e.g. Pinna et al 2007, Kleiger et al 1991, Sandercock et al 2004, Sinnreich et al 1998), meaning that our inferences from the present population are still widely applicable.

Conclusions

In conclusion, the present results demonstrate good reproducibility of HRV parameters during exercise, thereby enabling future studies involving exercise-related interventions. The importance of relative rather than absolute exercise intensity for the reliable determination of HRV parameters is illustrated and should be adhered to in future studies.
Acknowledgment

This work was supported by the Welsh Government (grant number: HE09 COL1029).

References

Aubert A E, Beckers F and Ramaekers D 2001 Short-term heart rate variability in young athletes J. Cardiol. 37 85–8
Eckberg D L 2009 Point:Counterpoint: Respiratory sinus arrhythmia is due to a central mechanism versus respiratory sinus arrhythmia is due to the baroreflex mechanism J. Appl. Physiol. 106 1740–2
Gladwell V F and Coote J H 2002 Heart rate at the onset of muscle contraction and during passive muscle stretch in humans: a role for mechanoreceptors J. Physiol. 540 1095–102
HRV reproducibility during exercise


Porges S W and Byrne E A 1992 Research methods for measurement of heart rate and respiration Biol. Psychol. 34 93–130


Taskforce 1996 Heart rate variability: standards of measurement, physiological interpretation, and clinical use Circulation 93 1043–65


Umetani K, Singer D, McCrathy R and Atkinson M 1998 Twenty-four hour time domain heart rate variability and heart rate: relations to age and gender over nine decades J. Am. Coll. Cardiol. 31 593–601

Whipp B J, Davis J A, Torres F and Wasserman K 1981 A test to determine parameters of aerobic function during exercise J. Appl. Physiol. 50 217–21

Ziegler D, Pióló R, Schabussen K, Lambeck H and Dammel H 1999 Normal ranges and reproducibility of statistical, geometric, frequency domain, and non-linear measures of 24 h heart rate variability Horm. Metab. Res. 31 672–9